

ORDER NO. 6096

UNITED STATES OF AMERICA
POSTAL REGULATORY COMMISSION
WASHINGTON, DC 20268-0001

Before Commissioners:

Michael Kubayanda, Chairman;
Ann C. Fisher, Vice Chairman;
Mark Acton;
Ashley E. Poling; and
Robert G. Taub

Periodic Reporting
(Proposal Six)

Docket No. RM2020-13

ORDER ON ANALYTICAL PRINCIPLES USED IN PERIODIC REPORTING
(PROPOSAL SIX)

(Issued January 26, 2022)

I. INTRODUCTION

On September 15, 2020, the Postal Service filed a petition pursuant to 39 C.F.R. § 3050.11, requesting that the Commission initiate a rulemaking proceeding to consider changes to analytical principles relating to periodic reports.¹ Proposal Six introduces a new methodology for estimating volume variabilities for certain mail processing cost

¹ Petition of the United States Postal Service for the Initiation of a Proceeding to Consider Proposed Changes in Analytical Principles (Proposal Six), September 15, 2020 (Petition). The Petition was accompanied by a study supporting its proposal. See A. Thomas Bozzo & Tim Huegerich, Analysis of Labor Variability for Automated Letter and Flat Sorting, Christensen Associates, September 15, 2020 (Variability Report). The Postal Service also filed a notice of filing of public and non-public materials relating to Proposal Six. Notice of Filing of USPS-RM2020-13-1 and USPS-RM2020-13-NP1 and Application for Nonpublic Treatment, September 15, 2020.

pools: Delivery Barcode Sorter (DBCS), Automated Flats Sorting Machine (AFSM) 100, and Flats Sequencing System (FSS). Petition, Proposal Six at 1. For the reasons discussed below, the Commission denies Proposal Six.

II. PROCEDURAL HISTORY

On September 23, 2020, the Commission issued a notice initiating this proceeding, providing for the submission of comments, and appointing a Public Representative.² Six Chairman's Information Requests (CHIRs) were issued.³ The Public Representative filed two motions for issuance of information requests, and subsequent CHIRs were issued based on his motions.⁴ The Postal Service provided clarifying information in its six responses.⁵ On November 24, 2020, the Commission

² Notice of Proposed Rulemaking on Analytical Principles Used in Periodic Reporting (Proposal Six), September 23, 2020 (Order No. 5694).

³ Chairman's Information Request No. 1, October 5, 2020 (CHIR No. 1); Chairman's Information Request No. 2 and Notice of Filing Under Seal, October 26, 2020 (CHIR No. 2); Chairman's Information Request No. 3 and Notice of Filing Under Seal, November 6, 2020 (CHIR No. 3); Chairman's Information Request No. 4, November 12, 2020 (CHIR No. 4); Chairman's Information Request No. 5, December 10, 2020 (CHIR No. 5); Chairman's Information Request No. 6, February 10, 2021 (CHIR No. 6). The Postal Service filed a motion to extend the deadline to respond to CHIR No. 5. See Motion of the United States Postal Service for an Extension of Time to Respond to Chairman's Information Request No. 5, December 11, 2020. The Commission granted the motion on December 14, 2020. Order Granting Motion for Extension of Time, December 14, 2020 (Order No. 5778).

⁴ Public Representative Notice of Filing Confidential Motion for Issuance of Information Request, November 2, 2020; Public Representative Motion for Issuance of Second Information Request, November 2, 2020. In response to the confidential motion, the Postal Service took no view on the merits of the motion but observed that the proposed questions could be posed and answered in public documents. Response of the United States Postal Service to Public Representative Motion for Issuance of an Information Request, November 2, 2020. See, e.g., CHIR No. 3.

⁵ Responses of the United States Postal Service to Questions 1-11 of Chairman's Information Request No. 1, October 14, 2020 (Response to CHIR No. 1); Responses of the United States Postal Service to Questions 1-8 of Chairman's Information Request No. 2, November 5, 2020 (Response to CHIR No. 2); Responses of the United States Postal Service to Questions 1-8 of Chairman's Information Request No. 3, November 13, 2020 (Response to CHIR No. 3); Responses of the United States Postal Service to Questions 1-4 of Chairman's Information Request No. 4, November 19, 2020 (Response to CHIR No. 4); Responses of the United States Postal Service to Questions 1-9 of Chairman's Information Request No. 5, January 5, 2021 (Response to CHIR No. 5); Responses of the United States Postal Service to Questions 1-7 of Chairman's Information Request No. 6, February 19, 2021 (Response to CHIR No. 6).

received comments regarding Proposal Six from MPA – The Association of Magazine Media, the Association for Postal Commerce, and the American Catalog Mailers Association (Joint Commenters) and the Public Representative.⁶ The Postal Service filed reply comments on December 8, 2020.⁷

III. BACKGROUND

Proposal Six introduces a methodology for estimating volume variabilities for certain mail processing cost pools: DBCS, AFSM 100, and FSS. Petition, Proposal Six at 1. The cost pools at issue involve labor expenses associated with the mail processing operations used for automated distribution of letters (DBCS) and flats (AFSM 100 and FSS). *Id.* at 2. The Postal Service states that accrued labor costs in these three cost pools totaled \$2.3 billion in FY 2019. *Id.* at 1.

The Postal Service asserts that the main factor determining labor requirements for sorting operations is the number of pieces inducted into the operation for processing, Total Pieces Fed (TPF) in the Management Operating Data System (MODS). *Id.* at 2; Variability Report at 7. In automated distribution operations, the actual number of handlings are directly counted by the sorting equipment and automatically transmitted from the equipment to the Web End-of-Run (WebEOR) system. Petition, Proposal Six

⁶ Comments of MPA – The Association of Magazine Media, the Association for Postal Commerce, and the American Catalog Mailers Association, November 24, 2020 (Joint Comments); Public Representative Comments on Proposal Six, November 24, 2020. The Public Representative submitted two revisions of his comments. See Public Representative Revised Comments on Proposal Six, November 25, 2020; Public Representative Rev.2 Comments, November 30, 2020 (PR Second Revised Comments). The Public Representative filed several motions related to his comments and supporting workpapers. See Public Representative Motion to File Erratum, November 25, 2020; Motion to File PR LR 1 Late, November 25, 2020; Public Representative Motion to File Erratum 2, November 27, 2020; Public Representative Motion to Revised RM2020-13, PR-LR-1, November 30, 2020. These motions are granted. Any reference to the PR Comments in this Order is to the PR Second Revised Comments.

⁷ Reply Comments of the United States Postal Service Regarding Proposal Six, December 8, 2020 (Postal Service Reply Comments). The Postal Service Reply Comments were accompanied by a report supporting the comments. See Reply Report of A. Thomas Bozzo in Response to Comments of the Public Representative, December 8, 2020 (Reply Variability Report). The Postal Service filed a motion for leave to file reply comments. See Motion of the United States Postal Service for Leave to File Reply Comments Regarding Proposal Six, December 8, 2020. The motion is granted.

at 2. MODS collects and aggregates piece handlings and runtime data through automated interfaces with WebEOR. *Id.* Labor usage or workhour data by operation are derived from time clock rings reported to MODS through the Time and Attendance Collection System (TACS). *Id.*; Variability Report at 15.

Currently, In-Office Cost System tallies are used to partition the mail processing cost pools into activities assumed to be 100 percent volume-variable, and other activities assumed to be non-volume-variable. *Id.* The basis for such determination was an assumption that mail processing costs should vary in proportion to the volume of mail or articles processed. See Variability Report at 4. For the operations that are the subject of this analysis, the associated mail processing costs were taken to be 99.1-percent volume-variable in FY 2019 under the accepted methodology. *Id.*

This methodology has been in use since Docket No. R71-1, and its origins predate the Postal Reorganization Act and the development of the automated mail processing technologies in this proposal.⁸ The Postal Service states that the Commission previously declined to adopt any empirical models for mail processing variability, citing data and econometric issues. *Id.* at 3.

“In Docket No. R71-1, the Commission considered empirical evidence, including a simple linear regression, of the relationship of clerk-mailhandler [workhours] to the ‘weighted volume of mail and special services.’” PRC Op. R97-1, ¶ 3010. Considering that “[t]he intercept of the linear regression was a small negative number, [which indicated] that the volume-variability of mail processing manhours was greater than 100 percent[,]. . . the Commission ‘classified’ the costs of ‘mail processing and related activities’ as ‘variable’ and [had] regarded them as 100 percent volume-variable.” *Id.*

⁸ Petition, Proposal Six at 2; see also Docket No. R97-1, Opinion and Recommended Decision Vol. 1, May 11, 1998, ¶ 3010 (PRC Op. R97-1) (citing Docket No. R71-1, Chief Examiner’s Initial Decision on Postal Rate and Fee Increases, February 3, 1972 (Docket No. R71-1 Decision)).

This finding also coincided with the operational judgement that “mail processing labor costs rise essentially in proportion to the volume of mail processed.”⁹

In Docket No. R97-1, the Postal Service’s witness Bradley proposed the first econometric model for estimating the variability of mail processing labor costs (Bradley Model). PRC Op. R97-1, ¶¶ 3001-3056. The Commission rejected the proposal due to a number of “disqualifying defects” in Bradley’s variability analysis. *Id.* at ¶ 3007. These defects were: too short (only 8 weeks) sample period, “error-ridden” and ineffectively cleaned MODS data, multiple model specification issues, and incorrect hypotheses regarding mail processing operations. *Id.* The Commission later noted that the data quality “was too poor to support a valid statistical model, that [the model] did not reflect an articulated economic theory, [and that the sample period was too short,] and that the resulting variabilities (76 percent) were so low as to be counterintuitive.”¹⁰

In Docket No. R2000-1, the Postal Service’s witness Bozzo, as well as a few other witnesses, proposed mail processing variability models, but the Commission did “not consider any of these econometric model results to be reliable.”¹¹ The Bozzo Model was based “on a novel view of how the Postal Service [deals] with volume changes over a typical rate cycle [and was] structured to capture primarily short-run volume effects.” *Id.* ¶ 3023. The Commission found “that the mail processing system that witness Bozzo [had modelled differed] substantially from the one observed in the real world.” *Id.* ¶ 3051. For example, piece handlings performed in an operation were assumed to be proportional to the volume of mail processed in that operation, although the econometric evidence and operational reasons suggested that “the number of

⁹ Docket No. R2000-1, Opinion and Recommended Decision Vol. 1, Revised November 30, 2000, ¶ 3020 (PRC Op. R2000-1).

¹⁰ Docket No. R2006-1, Opinion and Recommended Decision Vol. 1, February 26, 2007, ¶ 3005 (PRC Op. R2006-1) (citing PRC Op. R97-1, ¶ 3013).

¹¹ PRC Op. R2000-1, ¶ 3033. In addition to the Postal Service’s model (Bozzo Model), United Parcel Service, Inc.’s (UPS) witness Neels and witness Elliot representing the Periodicals mailers’ group, presented their models on the record. *Id.* ¶¶ 3032-3033.

handlings increase[d] faster than volume.” *Id.* ¶ 3033. In addition, the quality of the underlying data was no better than that upon which the initial Bradley model was based. *Id.* ¶ 3025. The Commission concluded that “obvious errors are common in MODS data [and] a substantial risk of ‘errors-in-variables’ bias remains after [applying] witness Bozzo’s data screens.” *Id.* ¶¶ 3026, 3028. The resulting variability estimates of 72.8 percent were even lower than the 81 percent variability estimates derived from Bradley Model and could not be “regarded as reliable.” *Id.* ¶¶ 3028, 3031, 3043.

In Docket No. R2005-1, the Commission found that the new mail processing variability analysis presented by the Postal Service witness Bozzo had several improvements over the models offered in previous dockets.¹² The Commission explained that an “instrumental variables” approach to mitigate significant bias in variability estimates for manual letter and flat sorting operations appears to validate the Commission’s concern that erroneous MODS data have a significant downward impact on variability estimates for mail processing generally. PRC Op. R2005-1, ¶ 4097. However, the Commission did not reach the merits of the presented variability analysis, noting that certain “properties of the MODS datasets...and their implications for econometric modeling...warrant[ed] further research.” *Id.* ¶ 4099.

In Docket No. R2006-1, the Commission evaluated three mail processing variability models – one developed by the Postal Service, one sponsored by the Office of the Consumer Advocate (OCA), and one developed by UPS. PRC Op. R2006-1, ¶ 3001. The Postal Service’s witness Bozzo provided some revisions to the model presented in Docket No. R2005-1 (Bozzo Updated Model). *Id.* ¶ 3010. The modeling approach recommended by OCA relied on a theory of production (developed by its witness Roberts) “that defines output as real volume, i.e., the number of unique pieces of mail processed in the plant.” *Id.* ¶ 3011. The modeling approach recommended by UPS was a multi-variate model developed by witness Neels (Neels Model) that

¹² Docket No. R2005-1, Opinion and Recommended Decision Vol. 1, November 1, 2005, ¶ 4097 (PRC Op. R2005-1).

estimated the response of workhours “in the entire plant to changes in [First Handled Pieces (FHP)] for each of the major shapes—letters, flats, and parcels.” *Id.* ¶ 3081. The Commission rejected all three models, stating that MODS data quality was a “formidable obstacle” to econometric modeling. *Id.* ¶ 3069. The Commission identified the lack of plant-specific data on true volume – “unique [Revenue Pieces and Weight (RPW)] pieces finalized to exit the mail processing system” as the most important obstacle. *Id.* ¶ 3083. In addition, the Commission found that both the Roberts Model and the Bozzo Updated Model failed to produce stable or intuitively reasonable results. *Id.* ¶¶ 3080, 3091. In regard to the Neels Model, the Commission concluded that although results appeared intuitively reasonable, they varied by the level of screening and required such “massive truncation of the available data that Neels, himself, d[id] not trust the results.” *Id.* ¶ 3091. Discussing the Bozzo Updated Model, the Commission specifically noted that the results of the variability analysis are very sensitive to changes in screening levels and that seasonal variation is not properly reflected in the model. *Id.* ¶¶ 3043, 3100.

In the current docket, the Postal Service asserts that “[s]everal factors merit re-examination of variabilities for automated letter and flat sorting[,]” including volume changes, the mature state of automated sorting equipment, reliability of automated counts of mailpiece handlings, and the availability of machine utilization data. Petition, Proposal Six at 4.

IV. SUMMARY OF PROPOSAL SIX

The Proposal Six methodology for estimating volume variabilities for automated mail processing cost pools is based on econometric analysis of workhour and workload data collected by the Postal Service on an ongoing basis. *Id.* at 1. Specifically, the estimation of the proposed variabilities employs monthly MODS datasets compiled into a multi-year panel dataset. *Id.* at 5. The variabilities are derived from a regression equation, where the natural logarithm of workhours is used as the dependent variable and the natural logarithms of TPF (current and lagged) as well as seasonal dummy

variables are used as explanatory variables. *Id.* The regression sample periods cover the most recent 4 fiscal years and would be rolled forward to allow for re-estimating the variabilities annually. *Id.* The variabilities estimated for the three cost pools during a FY 2016 to FY 2019 sample period are 0.976 for DBCS, 0.774 for AFSM 100, and 0.804 for FSS. *Id.* at 6.

The Postal Service states that the proposed methodology would permit re-estimation of the variabilities because the underlying data are produced in the course of Postal Service operations and are already included in the Annual Compliance Report. *Id.* at 1-2. The Postal Service concludes that the proposed methodology would reduce FY 2019 volume-variable labor costs for the three cost pools by 8.3 percent overall. *Id.* at 6. The Postal Service estimates relatively larger cost impacts for flat-shape products, and relatively smaller cost impacts for letter and parcel-shape products. *Id.* The Postal Service provides a table showing the effects of the proposed variabilities on product unit costs. *Id.* at 6-8. In a separate table filed under seal, the Postal Service shows the impacts of the proposal on individual Competitive products.¹³

V. COMMENTS

Joint Commenters and the Public Representative each submitted comments on Proposal Six.¹⁴ The Postal Service submitted reply comments to address the concerns of the Public Representative.¹⁵

A. Comments by Joint Commenters

Joint Commenters support Proposal Six, stating that its empirical estimation of mail processing labor variability “is an improvement over the current approach of simply assuming that these costs are essentially 100 percent variable.” Joint Comments at 2.

¹³ See Library Reference USPS-RM2020-13/NP1, September 15, 2020.

¹⁴ See Joint Comments; PR Second Revised Comments.

¹⁵ See Postal Service Reply Comments; Reply Variability Report.

Joint Commenters present an analysis aiming to illustrate that during the FY 2012 to FY 2019 period, there had been “much slower” declines in “automated flats sorting workhours” compared to the relative declines in TPF.¹⁶ Joint Commenters conclude that the observations corroborate the Postal Service’s finding that mail processing volume variability [for automated flats sorting equipment] is less than 100 percent. Joint Comments at 3.

Joint Commenters, however, argue the Postal Service “has not appropriately rightsized either its flats sorting workforce or infrastructure in response to declines in flats sorting workload” and urge the Commission to investigate “[the Postal Service’s] clear excess capacity and workhours in flats sorting operations.” *Id.* at 4, 6.

B. Public Representative Comments

The Public Representative strongly opposes the Proposal Six mail processing variability study, arguing that it differs little from previous models rejected by the Commission. PR Second Revised Comments at 1-2. The Public Representative maintains that an analysis of previous Commission dockets show that the Postal Service has not met “the burden of proof necessary to submit a new model of mail processing variability.” *Id.* at 16. The Public Representative asserts that in the past, the Commission laid down certain requirements for future mail processing studies. *Id.* at 9-10. The Public Representative also discusses previous Commission concerns arguing that the Postal Service failed to address them in its Proposal Six model. *Id.* at 17. Among these concerns is lack of the comprehensive measure of volumes exiting each mail processing plant; hidden errors in the analysis dataset due to the aggregation of uncleaned data and inappropriate application of 5 percent productivity screens to clean the already aggregated data. *Id.* at 17-18, 22-23. In regard to model specification, the Public Representative implies that the inclusion of the highly collinear lagged volume

¹⁶ *Id.* at 3; see also Excel file “RM2020-13 Workpapers.xlsx” filed with Joint Comments on November 24, 2020 (Joint Comments Workpapers).

variables in the regression models was not justified and its preference for the lagged model was not explicitly defended. *Id.* at 16, 24.

C. Postal Service Reply Comments

The Postal Service points out the key differences between the current and the 15-year old mail processing environment including the prevailing share of mature automatic sorting operations and the reliability of MODS volume data for the automatic mail processing operations. Reply Variability Report at 1-2. Responding to the Public Representative's concern regarding the volume measure used in Proposal Six, the Postal Service presents the analysis in support of "the use of TPF as the output measure determining sorting labor."¹⁷ Addressing the Public Representative's criticism of the econometric methodology underlying Proposal Six, the Postal Service defends the use of models with lags and seasonal variables noting that the "lagged output and seasonal dummy variables...are justified both on *a priori* considerations and with specification testing." Reply Variability Report at 9-10 (emphasis added) (internal citation omitted). In regard to multicollinearity (high collinearity between the current and lagged volume variables), the Postal Service states that the model "is not classically symptomatic of multicollinearity" because "many if not most of the [volume] coefficients are significant individually as well as jointly." *Id.* at 10. The Postal Service disagrees with the Public Representative's concern that the application of 5 percent productivity screens led to deletion of valid observations arguing that his analysis mistakenly used workhours instead of productivity to screen observations. *Id.* at 12-13. The Postal Service concludes that a substantial number of observations the Public Representative classifies as "wrongly deleted" are either unusable (e.g., because TPF, hours, or both are zero) or are not actually deleted from the regression samples. *Id.* at 13.

¹⁷ *Id.* at 10-11; see also Library Reference USPS-RM2020-13-5, December 8, 2020, folder "Analysis," Excel files "results_seasonal_fhp.xlsx;" "results_seasonal_fhp2.xlsx;" "results_seasonal_fhp_iv.xlsx."

VI. COMMISSION ANALYSIS

A. Overview

Pursuant to 39 C.F.R. § 3050.11(a), proposed changes to analytical principles are evaluated to ensure that they improve the quality, accuracy, or completeness of the data or analysis of data contained in the Postal Service's periodic reports. The Postal Service presents Proposal Six in order to "permit regular re-estimation and updating of the variabilities" for "the mail processing cost pools representing automated letter and flat sorting operations." Petition, Proposal Six at 1-2. The Commission finds that the applicable statutory requirements have not been satisfied because the Postal Service has not shown that the proposed revisions to the methodology for estimating volume variabilities would result in a significant improvement in the attribution of costs.

The Commission identifies a number of data issues related to Proposal Six. See Section VI.B. Although the quality of TPF data in MODS has generally improved, the workhours data may still be subject to measurement error and the Postal Service has not attempted to clean the data at the initial stage of developing the analysis dataset. See Sections VI.B.1 through VI.B.3. Previously, the Commission consistently stated that data cleaning should be done at the shift or daily level, but Proposal Six continues the practice of performing data cleaning after data aggregation to the monthly level, potentially allowing for errors to be masked by aggregation. See Section VI.B.3. In addition, to clean the already aggregated data, the Postal Service applies the productivity screening approach that has a number of flaws and is unacceptable in its current state. See Section VI.B.4.

The Commission also identifies a number of modeling and operational issues in Proposal Six. See Section VI.C. First, the Commission finds that the Postal Service does not properly justify its choice of lagged TPF variables that are included into the model to account for a delayed response of workhours to change in volume. See Section VI.C.2. Second, the Proposal Six econometric model likely excludes important control variables that are correlated with volume, which results in biased variability

estimates. See Section VI.C.3. Third, modeling the AFSM 100 operations at the aggregate level without controlling for some differences between subtypes of these operations is problematic. See Section V.D. The Commission also finds that the assumptions underlying the estimated variabilities are not well-justified, and minor changes in these arbitrary assumptions lead to the different variability estimates. See Sections VI.B. through VI.D. These problems preclude the Commission from both choosing between alternative estimates of mail processing variabilities and accepting the proposed methodology for their estimation. The Commission requests the Postal Service correct the identified problems before presenting a new mail processing variability study.

B. Data Issues

1. Overview

As briefly discussed in Section IV., the Postal Service's variability study employs the MODS data by month and facility to estimate labor variabilities for its three major automatic mail processing operations: DBCS, AFSM 100, and FSS. See Petition, Proposal Six at 5. In the variability models developed for each automatic operation, the Postal Service chooses workhours as the dependent variable and the TPF, a measure of the mail processing volume, as a primary explanatory variable. *Id.* In this section, the Commission summarizes its findings related to the quality of MODS data in general and the Proposal Six dataset, in particular. The Commission concludes that the quality of TPF data related to the automatic mail processing operations has generally improved since the Postal Service presented its last variability study in Docket No. R2006-1, which was more than 15 years ago. See Section VI.B.2. The Commission, however, finds that workhours data are still subject to measurement error. See Section VI.B.3. The Commission also has serious concerns about the Postal Service's productivity screening approach to data cleaning, which, in addition to being quite problematic in general, as discussed in Section VI.B.4., might aggravate the likely effect of the

measurement error in the workhours data. This is because the Postal Service applies the productivity screens to aggregated data. See Section VI.B.3. Most of these data cleaning issues the Commission identified in the previously rejected variability studies, and the Postal Service has not addressed them in Proposal Six.

2. The Quality of MODS TPF Data Related to Automatic Mail Processing has Generally Improved

As discussed in Section III., the poor quality of MODS data played a major role in the Commission's decisions to reject mail processing variability studies in the past. The Commission consistently noted that the MODS data were "too error ridden to produce estimates that [would be] sufficiently free of bias to be relied upon for ratemaking."¹⁸ In particular, the Commission was concerned about the quality of data for volume explanatory variables in the estimated variability models noting that these variables "must be substantially free of error [to avoid biasing] the estimated coefficient of that variable toward zero." PRC Op. R2006-1, ¶¶ 3030-3031 (footnote omitted). The Commission stated:

[T]he consequences that follow from using an explanatory variable measured with a substantial level of error can be severe. When fit to the sample, the model's estimated coefficients are likely to be biased and inconsistent. "Biased" means that the estimated coefficients aren't likely to be the true coefficients. "Inconsistent" means that the bias will remain even if the sample is large.

Id. ¶ 3030 referring to PRC Op. R97-1 at 82-83.

In Proposal Six, the Postal Service describes changes in the operational environment that have occurred in the 15 years since the last mail processing variability study. See Variability Report at 5-6; Reply Variability Report at 2-6. It correctly points out the mature state of automatic operations for flats and greatly diminished role of manual mail processing in the current system. See Variability Report at 5-6; Reply

¹⁸ PRC Op. R2006-1, ¶ 3029 referring to PRC Op. R2005-1, Appendix I at 29-33, 37, 52; PRC Op. R2000-1 at ¶¶ 3033, 3047; PRC Op. R97-1 at 81-84.

Variability Report at 1-2. In the previous mail processing variability studies, the Postal Service attempted to estimate variabilities for both manual and automated sorting operations, while the Proposal Six methodology is for automated operations only. Petition, Proposal Six at 3-4. In Docket No. R2006-1 and prior dockets “inaccurate weight-to-pieces conversion factors,” were used to determine FHP in each operation and also provided a basis for Total Pieces Handled (TPH) count for the manual operations. Reply Variability Report at 5; see *also* PRC Op. R2006-1, ¶¶ 3044, 3047.

The Commission previously noted that by using TPF, the Postal Service would be able to “obtain an accurate, unbiased estimate of volume variability” if the measurement of such variable is “substantially free of error.” *Id.* ¶¶ 3019, 3030. In the old variability studies, however, errors in the TPF data presented a serious issue in introducing bias into the variability models.¹⁹ In Proposal Six, the TPF data are machine-generated and their measurement is not subject to conversion. See Variability Report at 15, Reply Variability Report at 5. As the Postal Service explains, “[e]rrors-in-variables [bias from measurement error in TPF] is not an issue for Proposal Six because the MODS sorting volumes employed in the analysis are based on direct piece counts from the equipment that are automatically transmitted to MODS.” Reply Variability Report at 5; see *also* Response to CHIR No. 3, question 2.a. The Postal Service confirms: “Pieces of automated mail processing equipment generate end-of-run (EOR) files at the conclusion of each processing run. The EOR files provide data including equipment statistical data files and counts of...TPF by MODS operation and time, and are the sources of...TPF.”²⁰ The Commission therefore agrees with the Postal Service that the quality of MODS TPF data for automatic operations has improved in the last 15 years.

¹⁹ Docket No. R2006-1, Opinion and Recommended Decision Vol. 2, February 26, 2007, Appendix J ¶ 294 (PRC Op. R2006-1 Vol. 2).

²⁰ Response to CHIR No. 3, question 2.a.; see *also* Docket No. ACR2019, Library Reference USPS-FY19-7, December 27, 2019, folder “USPS-FY19-7,” PDF file “M-32 MODS Handbook.pdf,” at 8-9 (FY 2019 MODS Handbook).

In addition, to explore whether there is any evidence of the “explicit” errors in the MODS dataset that the Postal Service uses in its Proposal Six variability study, the Commission has performed an analysis similar to the one conducted by witness Neels in Docket No. R2006-1. PRC Op. R2006-1 Vol. 2, Appendix J ¶ 128. Table VI-1 below illustrates that the TPF data in the Proposal Six initial dataset appears relatively reliable compared to the analysis dataset witness Bozzo used in his variability model proposed in Docket No. R2006-1. For example, the incidences of obviously erroneous observations (such as those for which machine runtime is greater than zero, but workhours, TPF, or TPH are equal to zero) are much less common in the Proposal Six dataset than in the dataset used in the Docket No. R2006-1 variability study.

Table VI-1
Invalid Observations in the MODS Initial Datasets Used in Proposal Six and
Docket No. R2006-1 Variability Models

| Reasons for Observations Being invalid | Proposal Six ¹ (FY 2016-FY 2019) | | | Docket No. R2006-1 ² (FY 1999-FY 2005) | |
|--|--|----------|------|--|----------|
| | DBCS | AFSM 100 | FSS | DBCS | AFSM 100 |
| Runtime>0, but Workhours, TPF or TPH=0 | 0.2% | 0.1% | 0.0% | 11.4% | 3.1% |
| FHP>TPF | 0.0% | 0.0% | 0.0% | 13.5% | 20.7% |
| TPH>TPF | 0.0% | 0.0% | 0.0% | 0.2% | 1.6% |

Sources:

¹ Library Reference PRC-LR-RM2020-13/1, folder “MODS Data,” log file “analysis_seasonal_MODS.txt” and STATA program file “analysis_seasonal_mods.do.”

² PRC Op. R2006-1 Vol. 2, Appendix J at 49, Table J-3, lines 6, 11, 12.

Notes: TPF provides the number of all recorded mailpieces fed into the mail processing equipment. See FY 2019 MODS Handbook at 12. FHP is the number of first sorts of mailpieces. It is less than TPF by the number of rejected, reworked, or refed mailpieces. *Id.* TPH represents all the handlings of a mailpiece as it is processed, and it is less than TPF by the number of mailpieces rejected by mail processing equipment. See *id.* at 12, 17. As a result, by definition, $TPF \geq TPH \geq FHP$.

Based on the results of this analysis, the Commission concurs with the Postal Service that the measurement error in the TPF variable should not present any major issue for modeling unbiased labor variabilities. While TPF data are now machine-generated and appear to have improved, there are still some concerns regarding workhours data and the resulting reliability of the estimated variabilities as detailed in Section VI.3.

3. Concerns Remain Regarding Measurement Error in Workhours and Bias Resulting from Aggregation of Uncleaned Data

Contrary to TPF, which is a machine generated statistic, “workhours are derived from [employee] time clock rings reported to MODS through [TACS].” Variability Report at 15. The accuracy of workhours therefore depends on the extent to which employees are clocked into operation codes corresponding to their actual work activities and therefore workhours “[tend] to be more variable than processing equipment’s operating statistics.” *Id.*

In regard to the quality of workhours data, the Commission refers to the audit report by the United States Postal Service Office of the Inspector General (OIG) that reviewed Cost Segment 3 mail processing cost pool formation.²¹ The report found that adjustments to MODS data after the close of FY 2015 had effects on a number of MODS plants and Network Distribution Center cost pools, and the potentially misallocated costs were “only .01 percent of the total [FY 2015] mail processing costs.” CP-AR-17-007 at 8. The OIG, however, noted and the Commission agrees that “the ability to change MODS data after the close of the fiscal year without a tracking mechanism increases the risk of inaccurate cost data reports. It could also lead management and the Postal Regulatory Commission to rely on inaccurate information

²¹ See United States Postal Service, Office of Inspector General, Report No. CP-AR-17-007, Cost Segment 3 – Clerks and Mail Handlers, March 22, 2017, at 2, 8, available at <https://www.uspsaig.gov/sites/default/files/document-library-files/2017/CP-AR-17-007.pdf> (CP-AR-17-007).

when determining mail class and service attributable cost coverage and when setting postal prices.” *Id.* at 2.

The Postal Service notes that, in the econometric models for workhours, error-free workhour data are not required for avoidance of statistical bias and/or inconsistency because “[t]he measurement error variance is [just] a component of the residual variance.” Response to CHIR No. 1, question 3.a.; see *also* Variability Report at 15. The Commission, however, has previously expressed a concern about this argument stating:

If the measurement error in the dependent variable [workhours] is systematically related to the independent variable (output), then least-squares estimation methods can cause bias in the estimate of the output coefficient. A systematic relationship between misclocking and output might plausibly occur, for example, if [workhours] are underreported by workers whose productivity is evaluated on that basis relative to other workers.

PRC Op. R2006-1, ¶ 3033.

Despite these identified concerns, the Postal Service has not specifically studied the accuracy of MODS workhour data prior to using it in Proposal Six. See Response to CHIR No. 1, question 3.b. In Proposal Six, when constructing monthly observations, the Postal Service obtains the MODS raw data “by tour for each month” and then sums this uncleaned data “over the MODS tours.” Response to CHIR No. 3, question 2.g. The Postal Service asserts that workhours data are still relatively accurate at the high level of operational aggregation because “[e]rrors in workhour recording within cost pools or broader operation groups may not cancel to the same extent in a more disaggregated analysis, which may distort the calculation of accrued costs for disaggregated cost pools.” Response to CHIR No. 2, question 5.c.; see *also* Variability Report at 15. However, as the Public Representative points out, “[t]he Commission consistently rejected the Postal Service’s continual asides that aggregated data would wash out serious anomalies.” PR Second Revised Comments at 17-18. Thus, the Commission previously stated “MODS data are collected by shift, and rolled up by day, week, Accounting Period, and Quarter. Errors that would be explicit at the shift level tend to

be masked when rolled up into a daily count. The masking increases with the level of aggregation.” PRC Op. R2006-1, ¶ 3037.

The Commission affirms its previous conclusions and agrees with the Public Representative that the Postal Service has ignored previous Commission recommendations to clean data before their aggregation. PR Second Revised Comments at 17-18. The major problem here is that the Postal Service aggregates the collected data “without first attempting to remove gross errors or to distinguish missing observations from zero-valued observations.” *Id.* at 7. As a result, “incorrect or missing observations are combined with correct observations in the aggregated data sets.” PRC Op. R2005-1, Appendix I ¶ 1.

The Postal Service also recognizes data quality issues in the analysis datasets stating that the estimated variabilities for workhours are “somewhat sensitive to the inclusion of outliers with unusual values for labor productivity [TPF divided by workhours] in regressions using unscreened data.” Variability Report at 21. Because the TPF data is machine-generated and its quality has improved, and productivity is a function of TPF and workhours, unusual values for labor productivity are the result of unusual values of workhours. To remove the outliers from the analysis datasets, the Postal Service applies a productivity screening approach to the data already aggregated to the monthly, facility and operational group level. *Id.* As further illustrated by the analysis in Section VI.B.4., variabilities are sensitive to not only the inclusion of outliers in the unscreened dataset, but also a change in the cutoff values of the productivity screens. The Commission, therefore, concludes that it is impossible to determine whether the observed sensitivity is due to the MODS data errors masked by aggregation, unjustified productivity screen levels (and resulting cutoff values), or both these factors.

4. Major Flaws of the Productivity Screening Approach

To clean the analysis datasets for each of the three mail processing operations, the Postal Service applies a data screening approach that identifies non-“operationally-plausible” observations for which the “measured labor productivity” has extremely low or extremely high value. Variability Report at 21. Specifically, the Postal Service applies “the productivity cutoff values for the screen at the 5 percent tails,” (5 percent productivity screen) to remove observations (TPF at the facility and monthly level) that it considers outliers because their productivities are either “below the 5th percentile or above the 95th percentile of the distributions of site-month observations” for the operational group. *Id.*

The labor productivity that the Postal Service employs in its data screening criterion is a ratio between the mail processing volume (which is TPF in Proposal Six) and workhours. *Id.* at 21; *see also* Reply Variability Report at 14. The productivities are calculated as TPF divided by workhours for each facility and type of the automatic mail processing operation (DBCS, FSS and AFSM 100) using MODS data for the period between FY 2007 and FY 2019. Response to CHIR No. 5, question 7.c., Table 4. The Postal Service maintains that the proposed approach does not violate standard econometric practices, and explains that it allows the omission of observations for which productivity values are the result of “idiosyncratic errors or other factors not fully captured by the [Proposal Six] model[s].” Variability Report at 21; *see also* Reply Variability Report at 14.

The Commission agrees that prior to the econometric analysis, it is important to perform data cleaning, but as discussed in Section VI.B.3., it is critical to clean the data before aggregating it. In addition, the Commission identifies a number of problems specifically related to the productivity screening approach that the Postal Service applies in Proposal Six. The Commission finds that the proposed productivity screening is not a sufficiently rigorous approach to data cleaning. First, its underlying assumptions are questionable. Second, while the Postal Service fails to provide a clear justification

for any specific productivity cutoff values, the Commission analysis illustrates that the variability estimates are quite sensitive to changes in the productivity screen levels.

To determine the labor productivity cutoff values, the Postal Service explores the distribution of the productivity data for the Postal Accountability and Enhancement Act (PAEA) period from FY 2007 through FY 2019. See, e.g., Response to CHIR No. 5, question 7.c.iv. The productivities that are “below the 5th percentile or above the 95th percentile of the distributions of site-month observations” for the operational group are considered extreme values. Variability Report at 21. The Postal Service then removes the observations from the analysis dataset for the relevant operational group if the productivities for these observations fall into the range outlined above. One of the immediate problems here is that the variability analysis dataset for each operational group is developed using the MODS data for a different and much shorter time period of FY 2016 through FY 2019. *Id.* at 21. In order for the productivity screening approach to be sound, the operation-specific productivities identified as extreme for the 14-year sample period should be applicable to a shorter and more recent time period.

However, this proposition is not valid because, as illustrated by the Postal Service’s analysis, productivities have not been stable during the PAEA era, especially for flats operations. See *id.* at 16. The Postal Service states: “With generally declining scale, AFSM 100 productivity exhibits a more consistent downward trend. FSS productivity has also exhibited downward trends in productivity and average scale, though with a less consistent rate of decline.” *Id.* For DBCS operations, the productivity was not stable over the FY 2007 to FY 2019 span either. It declined initially to the lowest point in late FY 2009 (“when volume, workload, and scale all were declining...[, and then] productivity subsequently stabilized and increased...[because] facility consolidations drove an increase in average scale.” *Id.* The Postal Service also concedes that “declining productivity in flats operations effectively increased the strictness of the lower 5 percent productivity screen cutoff for the FY2016-2019 period compared to the entire FY2007-2019 period for which the cutoffs are computed.”

Response to CHIR No. 5, question 7.c.iv. The Postal Service has performed alternative calculations and computed the 5th/95th percentile productivity cutoff values based on the MODS data for the FY 2016 to FY 2019 sample period, the same that was used for variability analysis in Proposal Six. *Id.* Table 4. The variabilities estimated after application of these alternative productivity cutoff values are provided in Table VI-2 below.

In addition, the Commission analysis indicates, and the Postal Service confirms, that because productivity screens do not account for seasonal differences in productivities, the application of the productivity cutoffs results in the exclusion of relatively more observations from the months with higher productivities (such as November and December for DBCS) than from the months with lower productivities. See *id.* question 7.c.iv. The Postal Service has also computed the productivity cutoff values “by month to eliminate peak/off-peak season effects on the strictness of the boundaries.” *Id.* The re-estimated variabilities are also presented in Table VI-2 below. The Postal Service states that “[t]he regression analysis using the alternative screens shows *little effect* on the variabilities from these variations.” *Id.* (emphasis added).

The Commission finds that a similar conclusion would generally apply to the DBCS and FSS variabilities because the difference between the variabilities estimated under alternative screening methods is within 1 percentage point. See Table VI-2. For the AFSSM 100, however, the effect of applying alternative screens can be hardly characterized as having “little effect.” As illustrated in Table VI-2, the differences between variabilities estimated using the original and alternative screening methods are either close to or higher than 2 percentage points. The Commission concludes that the alternative productivity screening methods are conceptually more appropriate than the methods applied in Proposal Six. This is because they identify operation-specific extreme productivities using the productivity distribution over the same FY 2016 through FY 2019 time period as the Postal Service uses in the Proposal Six variability analysis.

Table VI-2
Workhour Variabilities Estimated under Alternative Screening Methods
(5 percent productivity screening level)

| Operation | DBCS | | AFSM 100 | | FSS | |
|---|-------------|----------------|-------------|----------------|-------------|----------------|
| Screening Method | Variability | Standard Error | Variability | Standard Error | Variability | Standard Error |
| Proposal Six Screen, FY 2007-FY 2019 data | 0.976 | 0.032 | 0.774 | 0.091 | 0.804 | 0.070 |
| Alternative screen, FY 2016-FY 2019 data¹ | 0.977 | 0.031 | 0.755 | 0.082 | 0.801 | 0.073 |
| Alternative screen, FY 2016-FY 2019 data,¹ monthly productivities² | 0.976 | 0.031 | 0.753 | 0.082 | 0.809 | 0.072 |

Sources: Response to CHIR No. 5, question 7, Table 4; see also Library Reference USPS-RM2020-13-6, January 5, 2021, folder "Question 7," file "analysis_seasonal_chir5q7c.txt;" Excel file "results_seasonal_chir5q7c.xlsx."

Notes: Proposal Six productivity cutoff values are originally calculated using data from the FY 2007 to FY 2019 time period and are constant across months. See Response to CHIR No. 5, question 7.c.iv.

¹ Variabilities are derived from the Proposal Six variability models, but the period for which productivity cutoff values are computed is changed to FY 2016 to FY 2019." All models are run on the same FY 2016 to FY 2019 dataset as in Proposal Six.

² Productivity cutoff values are computed by month.

The Commission, however, has reservations regarding the Postal Service's arbitrary choice of the 5 percent productivity screen. Discussing its method for choosing the cutoff values for productivity screening, the Postal Service notes that it was "not possible to set unambiguous [operation-specific productivity] cutoffs based on available information on machine throughput and staffing levels, particularly for AFSM 100 and FSS equipment subject to variable throughput and staffing levels." Response to CHIR No. 1, question 6.b. Therefore, to identify non-operationally plausible labor productivities, the Postal Service opts to use a method that would be "technically simple to apply and to update" and it would allow removing observations that are "at least...anomalous." Response to CHIR No. 5, question 6.a.; Reply Variability Report at 15.

The Postal Service further asserts that: “[e]ven at the 1% tails, it is perhaps notable that none of the [productivity] values...are clearly erroneous, particularly for DBCS operations. Observations for AFSM 100 operations less than half the median or FSS observations more than twice the median may both be regarded as at least being anomalous, and the 5% tails screen would exclude such values.” Reply Variability Report at 15. The Postal Service also notes that “observations just outside the [5th/95th percentile] cutoffs are unlikely to be materially erroneous observations.” Response to CHIR No. 5, question 7.c.iv. These statements support the Commission’s view that the Postal Service neither clearly determines what productivities are actually extreme, nor provides solid justification for its decision to choose the 5 percent productivity screen. The Postal Service appears amenable to alternate productivity screen cutoff values as it noted that it can use a “less restrictive” 1 percent tails screen instead of the 5 percent tails screen. Reply Variability Report at 15.

Also, as the Public Representative points out, the Commission previously stated that “deleting observations solely because they are unusual is not considered good econometric practice [because it]...is very possible that such ‘unusual’ observations contain the most information about the true relationship between cost and volume.” PR Second Revised Comments at 5 (citing PRC Op. R97-1 Vol. 2, Appendix F at 33). In response to the Public Representative criticizing the productivity screens for discarding valid data, the Postal Service clarifies that “the consequences of including erroneous data are generally more significant than those of not including all potentially valid data.”²² The Commission cannot accept the Postal Service’s argument as a basis for deleting observations in Proposal Six without providing a proper justification that they are erroneous. The Commission’s main concern is that the Postal Service does not have precise criteria for determining non-operationally plausible productivities and does

²² Reply Variability Report at 14. As to the Public Representative’s analysis that purports to show that Proposal Six procedures “wrongly” delete observations, the Commission agrees with the Postal Service that the Public Representative incorrectly identified “the set of observations covered by the screens” and therefore should be disregarded.” *Id.* at 12; see also PR Second Revised Comments at 23.

not have a strong rationale for a specific screening level (e.g., the 1 percent or the 5 percent tails screen) for computing the productivity cutoff value.

Considering the above, it is important to ensure that the variabilities are not sensitive to changes in productivity screen levels. To explore this issue, the Commission conducted an analysis comparing the workhour variabilities estimated using different screening levels (1, 2, 3, and 5 percent tails screen) and alternative screening methods for computing productivity cutoff values (which are described in the notes to Table VI-2 above). The analysis shows that under each screening method, the variabilities for flats mail processing operations (AFSM 100 and FSS) change substantially in response to a change in the screening level. See Table VI-3. Depending on the screening method, the variabilities estimated under the 1 percent tail screen or the 5 percent tail screen differ by 5 to 8 percentage points for both the AFSM 100, and the FSS operations. See Table VI-3. For flats operations, the higher the productivity cutoff value, the higher the estimated variability, and the lower (in most cases) the standard error.²³

The Postal Service does not strongly support the application of either the 1 percent tails screen or the 5 percent tails screen. The Commission's analysis reveals, however, that application of different screening levels leads to different variability estimates, without any specific choice appearing to be more correct than others. This casts a serious doubt on the validity of the provided screening method and reliability of the variabilities estimated in Proposal Six. Considering the identified flaws of the proposed productivity screening approach, the Commission concludes that it does not provide a reliable method for identifying erroneous observations.

²³ For details, see Library Reference PRC-LR-RM2020-13-1, folder "Alternative Screens," Excel file "results_seasonal_altscreen_PRC.xlsx."

Table VI-3
Workhour Variability Estimates under Alternative Methods and Levels of Productivity Screening

| Mail Processing Operation | Productivity Screening Method ¹ | Productivity Screening Levels (in percentage) | | | | Standard Error (range) |
|---------------------------|--|---|-------|-------|-------|------------------------|
| | | 1% | 2% | 3% | 5% | |
| DBCS | Proposal Six Screen, FY 2007-FY 2019 data | 0.992 | 0.980 | 0.979 | 0.976 | 0.032 - 0.035 |
| | Alternative Screen, FY 2016-FY 2019 data | 0.993 | 0.989 | 0.983 | 0.977 | 0.031 - 0.035 |
| | Alternative Screen, FY 2016-FY 2019 data, by month | 0.993 | 0.992 | 0.985 | 0.976 | 0.031 - 0.035 |
| AFSM 100 | Proposal Six Screen, FY 2007-FY 2019 data | 0.722 | 0.739 | 0.753 | 0.774 | 0.088 - 0.095 |
| | Alternative Screen, FY 2016-FY 2019 data | 0.673 | 0.714 | 0.733 | 0.755 | 0.082 - 0.102 |
| | Alternative Screen, FY 2016-FY 2019 data, by month | 0.672 | 0.711 | 0.731 | 0.753 | 0.082 - 0.103 |
| FSS | Proposal Six Screen, FY 2013-FY 2019 ² | 0.750 | 0.740 | 0.766 | 0.804 | 0.070 - 0.092 |
| | Alternative Screen, FY 2016-FY 2019 data | 0.744 | 0.739 | 0.760 | 0.801 | 0.073 - 0.094 |
| | Alternative Screen, FY 2016-FY 2019 data, by month | 0.729 | 0.747 | 0.763 | 0.809 | 0.072 - 0.100 |

Notes:

¹ The difference between three alternative screening methods is described in the notes to Table VI-2. The variabilities estimated using the same productivity screening method and level as in Proposal Six are highlighted in "green."

² For FSS that was deployed mainly in FY 2010 to FY 2011, the Postal Service used a shorter time period for productivity screening than for AFSM 100 and DBCS. See Variability Report at 11-12; Library Reference USPS-RM2020-13-1, folder "Analysis," Excel file "results_seasonal.xlsx," tab "FSS hrs."

Sources: For Proposal Six variability estimates, see Library Reference USPS-RM2020-13-1, folder "Analysis," Excel file "results_seasonal.xlsx." For alternative variability estimates, see Library Reference USPS-RM2020-13-6, folder "Question 7," Excel file "results_seasonal_chir5q7c.xlsx;" Library Reference PRC-LR-RM2020-13-1, folder "Alternative Screens," Excel file "results_seasonal_altscreen_PRC.xlsx," tab "Table VI-3."

C. Modeling Issues

1. Overview

As it discussed in Section VI.B.2., the Commission acknowledges the general improvements in the quality of the MODS data related to the automatic mail processing operations. This is a significant improvement of Proposal Six compared to the previously rejected variability studies where poor data quality was the main obstacle to econometric modeling. See Section III. Despite these improvements, the Commission identifies a number of issues directly related to the Proposal Six model. First, the Postal Service has not clearly justified its choice of the lagged explanatory variables. See Section VI.C.2. Second, the Commission suggests that the Proposal Six model omits important control variables, which likely results in biased variability estimates.

2. Problematic Choice of the Lagged Explanatory Variables

In its variability study, the Postal Service applies the model with lagged explanatory variables.²⁴ Specifically, each Proposal Six econometric model includes the current and two lagged TPF variables (the first month and the twelfth month lags) as explanatory variables. The Public Representative argues that “[t]he Postal Service [has] decided to include lagged volumes variables with little justification other than that there might be some additional information included in [the] model results.” PR Second Revised Comments at 25. As discussed below, the Commission agrees with the Public Representative regarding the Postal Service’s use of the lagged TPF variables without justification. The Public Representative also highlights the issue of “nearly perfect [correlation between the] current and lagged volume variables for each of the three machine types.”²⁵ The Public Representative further states that “a test of joint

²⁴ For more information about the models with lagged variables, see, e.g., Michael D. Intriligator, *Econometric Models, Techniques, & Applications*, 176-86 (1978).

²⁵ *Id.* (citing Library Reference USPS-RM2020-13-3, November 5, 2020, folder “Question 4,” log file “analysis_seasonal_chir2vif.txt”).

significance of variables [that] are highly collinear almost guarantee [that they] are jointly significant...because regression procedures cannot capture the shared effects” of information. PR Second Revised Comments at 26.

The Postal Service explains that it has chosen a model with lagged TPF variables because it captures “a lagged effect of TPF on...workhours,” which “allows for adjustment processes of workhours with respect to workloads over longer time scales.” Variability Report at 20. The Postal Service further states that “including lagged volumes helps distinguish whether trend differences between workloads and workhours are due to longer-term labor adjustment processes (i.e., lagged volume variability) versus non-unitary labor elasticities (i.e., volume variability factors not equaling 100 percent).” Response to CHIR No. 1, question 2.a. The overall labor variability is calculated as the sum of the current and two lagged TPF coefficients, respectively. Variability Report at 21; *see also* Response to CHIR No. 2, question 2.

For purposes of the variability estimates in Proposal Six, the Commission generally accepts the Postal Service’s selection of the distributed lag regression model, but questions the choice of the lagged TPF variables. Justifying its two specific lags, the Postal Service notes that the twelfth lag reflects the effects of the same-period-last-year (SPLY) operations on the current-period operations, while “the first lag is intended to capture shorter-term labor inflexibilities.” Response to CHIR No. 1, question 2.a. However, it appears unclear whether by including the first and the twelfth lags into the model, the Postal Service fully addresses these operational reasons and also ensures the best fit of the model. The Commission’s concern is that the variability model with the first and twelfth lags might fail to capture some additional “lagged effect of TPF on runtime and workhours” that it is intended to capture. Variability Report at 20. Thus, the Postal Service does not explain why the model that includes the first lag only (but does not include the second lag in addition to the first lag) fully captures “shorter-term labor inflexibilities.” Response to CHIR No. 1, question 2.a. Another concern is related to the inclusion of the twelfth lag to capture the effects of the SPLY operations while

excluding (and not even testing) other lags from the second to the eleventh. Measuring the effect of the conditions that prevailed last year on the current year's cost attribution may be an attempt to capture the slow adjustments of workhours to volumes. However, the need to account for such a long period of adjustment is far from being obvious and the Postal Service does not substantiate that need.

The inclusion of the specific two lags, the first and the twelfth lags only, into all three variability models is problematic because other lags could also be statistically significant if included into the variability model. As the Postal Service confirms, it has not econometrically tested the Proposal Six variability models with any lags besides the first and the twelfth. Response to CHIR No. 1, question 2.b. However, the Commission's analysis shows that such testing would be appropriate. Table VI-4 presents some preliminary results of model testing (*i.e.*, without the twelfth lag or with the second lag included into the model in addition to the first and twelfth lag). These results are discussed below by each type of mail processing operation.

Table VI-4
Preliminary Lag Testing in the Proposal Six Models

| Type of Operation | Lags Included | Estimated Variability | Standard Error | Are Lags Individually Significant? | Are Lags Jointly Significant? |
|-------------------|--|-----------------------|----------------|------------------------------------|-------------------------------|
| DBCS | 1 st and 12 th | 0.976 | 0.032 | Yes | Yes |
| | 1 st , 2 nd , and 12 th | 0.990 | 0.036 | No (the 12 th is not) | Yes |
| AFSM 100 | 1 st and 12 th | 0.774 | 0.091 | No (the 12 th is not) | No |
| | 1 st only | 0.818 | 0.083 | No | N/A |
| | 1 st and 2 nd | 0.776 | 0.066 | No | No |
| FSS | 1 st and 12 th | 0.804 | 0.070 | Yes | Yes |
| | 1 st , 2 nd , and 12 th | 0.828 | 0.071 | Yes | Yes |

Notes: Lags are tested for statistical significance at a 5 percent level using a t-test. To test for joint significance of lags, the F-test was applied. The variabilities estimated in Proposal Six are highlighted in “green.”

Sources: Data are from Library Reference USPS-RM2020-13-1, folder “Analysis,” Excel files “results_seasonal.xlsx;” “analysis_lag_seasonal_tests.txt.” For complete results, see Library Reference PRC-LR-RM2020-13-1, folder “Lag Testing,” log files “analysis_lag_seasonal_tests_PRC2.txt” (for DBCS and FSS testing); “analysis_lag_seasonal_tests_PRC3.txt” and “analysis_lag_seasonal_tests_PRC” (for AFSM testing).

After inclusion of the second lag in the DBCS model, the lags continue to be jointly significant, but the twelfth lag becomes individually insignificant. This result indicates that the twelfth lag most likely does not belong in the DBCS model, but also that it might be worth testing other lags (e.g., third lag, fourth lag, etc.) for inclusion into the model. The Commission’s preliminary conclusion is that for the DBCS model, first and second lags may sufficiently explain any relevant “adjustment processes of workhours with respect to workloads.” Variability Report at 20.

For the FSS model in Proposal Six, the lagged variables are jointly significant, and the first and twelfth lags are also statistically significant individually. With the second lag added, the lagged variables continue to be jointly significant, and each lag (namely, first, second, twelfth lags) is also statistically significant individually. The

difference between the estimated variabilities is, however, 3 percentage points. This difference is too large to be ignored, and is not consistent with the Postal Service's argument that "the Proposal Six results are robust to alternative specifications." Reply Variability Report at 10. The Commission suggests that the Postal Service should have expanded its lag testing and at least included the second lag and, potentially, other lags in the FSS model. It appears appropriate to continue testing for the individual and joint significance of lagged variables in the FSS model until the maximum number of statistically significant lags is obtained.

The situation is different for the AFSM 100 model. In the Proposal Six AFSM 100 model, the lags are not jointly significant, and the twelfth lag is individually insignificant.²⁶ The Postal Service explains that it adopted the AFSM 100 variability model despite the failure of statistical tests "because the null hypothesis that the coefficient on the first lag of TPF is zero was rejected at standard significance levels (p-value less than 0.01), and the null hypothesis for the joint [significance] test could be rejected in other sample periods (e.g., FY2015-2019) at standard significance levels (p-value less than 0.05)." Response to CHIR No. 6, question 4.b. As an additional justification for adopting the AFSM 100 model with the same lags as in the other models, the Postal Service states that it "would minimize the need to pretest the AFSM 100 model specification during the course of updating the models to incorporate new data." *Id.*

The Commission does not find the provided response convincing for multiple reasons. Contrary to its statement that the choice of lags is justified by "standard specification test statistics," the Postal Service does not adhere to commonly acknowledged econometric practice. Reply Variability Report at 10. For example, the Postal Service appears to reject the implications of standard specification test statistics

²⁶ The Postal Service states that the test for joint significance does "not reject the null hypothesis at standard significance levels (p-value 0.13)." Variability Report at 24 n.10; see also Library Reference USPS-RM2020-13-1, folder "Analysis," log file "analysis_lag_seasonal_tests.txt."

from the Proposal Six dataset (in regard to the specification test concluding that the twelfth lag was individually insignificant and lags were jointly insignificant) just because the first lag is still individually significant or lags are jointly significant when the model is run on data for “*other sample periods.*” Response to CHIR No. 6, question 4.b. (emphasis added). Preliminary testing of lags (*i.e.*, removing the twelfth lag from the model or adding the first lag instead of the twelfth lag) shows that lags are neither jointly or individually statistically significant. See Table VI-4. The Commission’s interim conclusion is that either both lagged TPF variables do not have any meaningful explanatory power in the Proposal Six AFSM 100 model, or separate variability models should be set up for different subtypes of the AFSM 100 operations. The last proposition is discussed in Section VI.D.

The Postal Service argues that it uses the minimum number of lags as a remedy to the multicollinearity problem identified by the Public Representative, but, as described below, moderates that argument when describing the multicollinearity issue. Reply Variability Report at 10. The Postal Service states that the main reason why it has not examined models with the full set of lags is the likelihood that “such a specification would encounter multicollinearity issues leading to statistically unreliable estimates of the coefficients on many or most of the lagged TPF variables.” Response to CHIR No. 4, question 3.d. The Commission agrees with the Postal Service that by adding additional lagged variables it could aggravate a multicollinearity problem pointed out by the Public Representative. However, the Postal Service provides other arguments that appear to dismiss the concern expressed in the statements quoted above. The Postal Service notes: “multicollinearity...does not appear to be a major problem for the analysis [and] the presence (or absence) of multicollinearity does not affect the statistical bias or consistency of the estimated coefficients, but would tend to increase the standard errors of the affected coefficient estimates.” Response to CHIR No. 2, question 4. In addition, when addressing the Public Representative’s concerns regarding nearly perfect collinearity among current and lagged variables, the Postal Service states that in Proposal Six, “many if not most of the coefficients are significant

individually as well as jointly [and that] is not classically symptomatic of multicollinearity.” Reply Variability Report at 10. The Commission concludes that if the coefficients for the current and lagged variables are individually and jointly significant, the Postal Service should not be too concerned about multicollinearity in the model.

In addition, the Commission finds that there is no applicable explanation that the short-term labor inflexibilities are fully captured by the first lag. The results of the analysis presented in Table VI-4 do not support the Postal Service’s proposition; the statistical significance of the second lag in the FSS variability model seem to prove that the workhours would continue to respond to a change in mail processing volumes after 2 months. Following the Postal Service’s explanation that the operational basis for lags besides the first and the twelfth lags would be to capture some additional “lagged effect of TPF on runtime and workhours,” the Commission finds the Postal Service’s methodology for determining lags in the Proposal Six model overemphasizes the expediency of updating the models year-over-year over the applicability and fit of the model. Variability Report at 20. While the Postal Service argues that “including the statistically insignificant lag does not lead to statistical bias or inconsistency,” Table VI-4 illustrates that a change in the number and/or types of lags results in different variability estimates, which would produce different volume variable costs and cost impacts. Response to CHIR No. 6, question 4.b.

The Commission concludes that the Postal Service does not provide sufficient support for the use of the specific lagged variables in the Proposal Six variability models. If the Postal Service makes a decision to apply the model with lagged variabilities in future variability studies, the Commission recommends that the Postal Service econometrically test different sets of lags before presenting a new model.²⁷

²⁷ To determine the number and types of lagged variables for the model, it might be useful to explore different methods, such as those suggested by Almon or Koyck. See J. Johnston, *Econometric Methods*, 294-300 (1960).

3. The Model is Likely Missing Important Control Variables

a. Advantages and General Limitations of the Fixed-Effects Model

The Proposal Six econometric model employs volume (namely, TPF) as the main explanatory variable. Variability Report at 21. To account for seasonality and to ensure that the “purely seasonal effects on workhours” are not “inappropriately” captured in the estimated variabilities, the model also includes monthly dummy variables. *Id.* at 20-21.

Since there might be other factors that also affect mail processing labor variabilities, the Postal Service uses the “plant” (or “facility”) level fixed-effects “to account for unobserved non-volume heterogeneity among facilities.”²⁸ The Postal Service includes these facility-specific fixed-effects into the model, as it has done in previous dockets, to account for “unobserved non-volume heterogeneity among facilities” and to avoid the problem of “biased and inconsistent [variability] estimates.”²⁹

The Postal Service explains that “technological parameters, management considerations affecting staffing levels locally, and specific plants’ processing network roles” may differ systematically across facilities. Variability Report at 20; *see also* Response to CHIR No. 3, question 3.b. These factors would all theoretically correlate with both volume and workhours and, as the Commission previously stated, “would induce a degree of omitted variables bias [in the model] if left out.”³⁰ The Postal Service

²⁸ *Id.* at 20. For a discussion about the use of fixed-effects as an estimation method in a regression model with unobserved (latent) variable(s), *see, e.g.,* Cheng Hsiao, *Analysis of Panel Data*, 27-30 (2003).

²⁹ Variability Report at 20. “The estimated variabilities are ‘biased’ when they are unlikely to be the true values, and the variabilities are ‘inconsistent’ when the bias remain even if the sample is large. *See* Badi H. Baltagi, *Econometrics*, 16-20 (2008). For discussion of biased and inconsistent mail processing variability estimates, *see* PRC Op. R2006-1, ¶ 3030; PRC Op. R2000-1 Vol. 2, Appendix F at 47; PRC Op. R97-1, ¶ 3040.

³⁰ *See* PRC Op. R2006-1, ¶ 3099. Omitted variable bias occurs when the omitted variable, first, correlates with an explanatory variable of interest (such as mail processing volume) and, second, determines the dependent variable (workhours) in the model. *See, e.g.,* Stock and Watson, *Introduction to Econometrics*, 187 (2007) (Stock and Watson, 2007); *see also* PRC Op. R2000-1 Vol. 2, Appendix F at 47.

asserts that because it cannot observe or quantify the key underlying factors that drive volume and workhour differences across facilities, “it is unlikely that the use of the fixed effects model can be avoided.” Response to CHIR No. 5, question 5.b.

To justify its use of a fixed-effects instead of a random-effects model, the Postal Service has conducted the Hausman test on each of its operation models. Response to CHIR No. 3, questions 3.a., 3.c. The resulting statistics reject[ed] the random-effects model for all three operation groups” in Proposal Six.³¹ The STATA software used by the Postal Service has also automatically tested the fixed-effects, and they are jointly significant.³² The Commission agrees that, generally speaking, the fixed-effects model may be superior to a model without fixed-effects. However, the Postal Service has not presented any compelling evidence that the provided variability estimates are reliable. For the reasons discussed below in Section VI.C.4.b., the Commission concludes that the Proposal Six fixed-effects model likely cannot account for some important time-variant, non-volume differences across facilities that are correlated with volume. This would lead to the omitted variable bias in the variability model, which, in turn, would cause variability estimates to be both biased and inconsistent. In addition, the Commission previously expressed a concern regarding limitations of employing quarterly dummies in the mail processing variability models, and Proposal Six does not address this concern. See PRC Op. R2006-1, ¶ 3100. As discussed in Section VI.C.4.c., it might be appropriate to replace monthly dummy variables with other variables to control specifically for non-volume seasonal effects.

³¹ *Id.* question 3.c. For the assumptions of the random-effects model, and implications of violations of these assumptions, see, e.g., Cheng Hsiao, *Analysis of Panel Data*, at 33 (1986).

³² See Library Reference USPS-RM2020-13-1, folder “Analysis,” log file “analysis_seasonal.txt.”

b. Considerations Regarding Additional Control Variables

The facility fixed-effects, as the term suggests, can only capture factors that are constant over time. PRC Op. R97-1, Appendix F at 10; see *also* Response to CHIR No. 6, question 5.c. The Postal Service has not provided sufficient proof that differences between plants that the fixed effects are meant to capture in the Proposal Six model are actually constant over time, and there is evidence that these differences may fluctuate. The Postal Service admits that different factors, such as management effectiveness, facility layouts, and local demographics, all affect workhours and correlate with volume. See Response to CHIR No. 5, question 5.a. The Postal Service has provided no evidence to substantiate its claims that the operational environment was fixed throughout a 4-year analysis sample period with regards to these factors. Discussing obstacles for quantifying such factors, the Postal Service states that “[e]ffects of facility layouts on workhours [, for example,] are not likely to be simple parametric functions of available statistics such as total facility square footage.” *Id.* question 5.b. The Postal Service, however, provides neither a basis for these statements nor results of the actual testing of this hypothesis. Similarly, even for “slow-changing factors” such as delivery network statistics or “largely constant” facility characteristics such as facility type that may determine workhours and relate to volume, the Commission emphasizes that *any* change in such factors would not be captured by the facility fixed-effects. *Id.* questions 5.b., 9.f. (emphasis added). While the data required to account for time-varying factors in the model may be not fully available at this time, the Commission recommends the Postal Service perform an in-depth exploration of possible ways to collect data on plant-level and time-varying factors in the future mail processing variability studies.

The Postal Service explains that “a partial motivation for employing a relatively short time period for the regression sample periods” is to limit “the amount of time variation in [the listed above] factors such as management quality, facility layouts, or local demographics.” Response to CHIR No. 3, question 3.b. To justify its choice of the FY 2016 through FY 2019 time period for the Proposal Six variability study, the Postal

Service emphasizes that this 4-year “period features a relatively fixed operating environment including technology mix.” Variability Report at 21.

The Commission does not find sufficient justification for the Postal Service’s choice of a 4-year sample period. As illustrated in Table VI-5, the estimation of the Proposal Six model using a 3-year subsample results in quite different variabilities than those estimated on a full 4-year sample. This difference is especially notable for DBCS and AFSM 100 operations (approximately 6 and 8 percentage points, respectively). See Table VI-5. These different variability estimates would result in materially different cost impacts for AFSM 100 and DBCS operations.

Table VI-5
Workhour Variability Estimates for Alternative Sample Periods (5 percent productivity screen)

| Operation | Time Period | Variability | Standard Error |
|------------------|---|--------------------|-----------------------|
| DBCS | Proposal Six, 4-year period, FY 2016-FY 2019 ¹ | 0.976 | 0.032 |
| | Alternative 3-year period, FY 2017-FY 2019 ² | 0.925 | 0.025 |
| | Alternative 5-year period, FY 2015-FY 2019 ² | 0.990 | 0.029 |
| | | | |
| AFSM 100 | Proposal Six, FY 2016-FY 2019 ¹ | 0.774 | 0.091 |
| | Alternative 3-year period, FY 2017-FY 2019 ² | 0.850 | 0.085 |
| | Alternative 5-year period, FY 2015-FY 2019 ² | 0.731 | 0.072 |
| | | | |
| FSS | Proposal Six, FY 2016-FY 2019 ¹ | 0.804 | 0.070 |
| | Alternative 3-year period, FY 2017-FY 2019 ² | 0.789 | 0.080 |
| | Alternative 5-year period, FY 2015-FY 2019 ² | 0.839 | 0.078 |

Sources:

¹ Library Reference USPS-RM2020-13-1, folder "Analysis", Excel file "results_seasonal.xlsx."

² Library Reference PRC-LR-RM2020-13-1, folder "Alternative Period," Excel file "results_seasonal_time_period.xlsx;" see *a/s/o* CHIR No. 6, question 7.a., Table 2.

The Postal Service states that “the effect of [moving from the 4-year period to the 3-year] period is to reduce the sample sizes for the regressions by approximately 25 percent for each operation.” Response to CHIR No. 6, question 7.c. The Postal Service still suggests that a three-year time period “may represent a reasonable alternative to Proposal Six – and a superior alternative to the existing variability assumptions – if the Commission were to weigh factors favoring a shorter time period more heavily.” *Id.* question 6.a. In addition, the Postal Service has submitted a version of its Proposal Six model using a 5-year sample for FY 2015 through FY 2019.³³ The Postal Service notes that there has been an absence of significant facility consolidation since FY 2015. Variability Report at 13. It appears the Postal Service suggests that the FY 2015 to FY 2019 time period could also be a reasonable alternative to the FY 2016 to FY 2019 sample period used in Proposal Six. However, although this time period would provide a larger sample size, the estimated variabilities are quite different from both the Proposal Six variabilities and the FY 2017 to FY 2019 variabilities for all three automatic operations. For FSS operations that were still in deployment in FY 2012, and for which a longer time period may not be as suitable, the Commission observes that the variabilities that the Postal Service reports in the rolling regression analysis of both 4-year and 5-year sample periods range from about 0.9 to around 0.8.³⁴

Considering a potential cumulative impact of this sensitivity to small changes in the sample period and other Proposal Six issues discussed in this Order (including, but not limited to, productivity screening approach to the data cleaning and choice of two specific lagged explanatory variables for the distributed lags model), the Commission

³³ See Library Reference USPS-RM2020-13-1, folder “Analysis,” Excel file “results_seasonal.xlsx.”

³⁴ See Variability Report at 24-26. In its rolling regression analysis, the Postal Service estimates variabilities using various 48-month and 60-month sample periods with the purpose to “investigate the stability of the [variabilities] within the overall FY2007-FY2019 period.” *Id.* at 24. The Postal Service opted to “examine rolling 60-month windows to investigate the effects of extending the sample period.” *Id.*; see also Library Reference RM2020-13-1, folder “Workbooks,” Excel file “Figs 11-12 rolling results.xlsx.”

does not find the Proposal Six methodology reliable. The Commission cannot affirm, based on the record in this case, any particular time length for the sample period and recommends that in future variability studies the Postal Service perform testing of alternative time periods to ensure consistency of the estimates. The Commission suggests that greater consistency would be achieved if the data trimming were applied to the non-aggregated MODS data sample taken over the same time period as that used for econometric analysis. See Section VI.B.3. and VI.B.4. Adding annual dummy variables into the model could also potentially resolve the consistency issue. Facility size is a variable that the Commission has previously identified as most likely not accurately captured in the Postal Service's labor variability models for AFSM 100 operations. See PRC Op. R2006-1 Vol. 2, Appendix J ¶ 243. In the past, the Commission found that "the estimated variabilities for the majority of the modeled operations [were] statistically significantly different depending on the size of the plant being modeled." *Id.* In the previous dockets, the Commission asked the Postal Service to complete this analysis to gauge the ability of its models to pool small plants and large plants. *Id.* The different results for small plants and large plants caused the Commission to conclude that "[a]nother reason that the Commission doubts that...variability estimates [in the previous variability studies] are reliable is that they are not consistent across subsamples." *Id.* The Commission has conducted a similar analysis for Proposal Six, by estimating its variability models separately for small, medium, and large facilities. Thus, for AFSM 100 operations, the variabilities estimated separately for small, medium, and large facilities are statistically significant, but differ substantially. This difference between variability estimates is statistically significant. See Table VI-6. Furthermore, the medium and large variabilities are much lower than the corresponding Proposal Six variability of 0.774 estimated for the AFSM 100 operations. The Commission urges the Postal Service investigate the reason for such a steep drop in variabilities for medium and large facilities in the future variability studies.

Table VI-6
Workhour Variabilities Estimated Separately by Facility Size

| Operation | Facility Size ¹ | Variability | Standard Error | Lower Bound of | Upper Bound of | Does Variability Fall Within the 95% Confidence Interval of the Other Two Facility Sizes? |
|-----------|----------------------------|-------------|----------------|-------------------------|----------------|---|
| | | | | 95% confidence interval | | |
| | | (1) | (2) | (3) | (4) | (5) |
| DBCS | Small | 1.019 | 0.081 | 0.860 | 1.177 | Yes |
| | Medium | 0.977 | 0.057 | 0.866 | 1.088 | Yes |
| | Large | 0.913 | 0.061 | 0.794 | 1.032 | Yes |
| | | | | | | |
| AFSM 100 | Small | 0.901 | 0.075 | 0.755 | 1.047 | No |
| | Medium | 0.557 | 0.087 | 0.386 | 0.728 | No |
| | Large | 0.599 | 0.080 | 0.441 | 0.756 | No |
| | | | | | | |
| FSS | Small | 0.761 | 0.135 | 0.498 | 1.025 | Yes |
| | Medium | 0.759 | 0.148 | 0.468 | 1.050 | Yes |
| | Large | 0.774 | 0.050 | 0.676 | 0.871 | Yes |

Sources: Library Reference PRC-LR-RM2020-13-1, folder "Facility Size," Excel file "results_seasonal_facility_size.xlsx," tab "Table VI-6," STATA Program file "analysis_seasonal_facility_size.do," and log file "analysis_seasonal_facility_size.txt."

Notes: In all models, the Commission applies 5 percent productivity screen and uses data for FY 2016 to FY 2019, as it is done by the Postal Service in Proposal Six. The formula for column (3) is $(3)=(1)-1.96*(2)$ and for column (4) is $(4)=(1)+1.96*(2)$.

¹ "Small" represents a model that includes observations (TPF at the facility and monthly level) below the 33rd percentile of the TPF within the operation group. "Medium" represents a model with observations between the 33rd and 66th percentile of the TPF within the operating group. "Large" represents a model with observations above the 66th percentile of the TPF within the operating group.

In regard to other explanatory variables to potentially use in a mail processing variability model, the Postal Service states that it "considered adding network and/or capital variables to the models, somewhat similar to the models employed in Docket No. R2006-1." Response to CHIR No. 1, question 2.c. The Postal Service, however, has rejected these variables for the following reasons: (1) high overall R-squared values in

the Proposal Six models,³⁵ (2) “generally small and statistically insignificant” coefficients of the network and/or capital variables in the previous variability models, and (3) high collinearity of explicit network variables with facility fixed-effects in Proposal Six. Response to CHIR No. 1, question 2.c. The Postal Service’s witness Bozzo previously stated that “[w]hether the model specifications can be modified to quantify the effects of the network on mail processing labor cost with low standard errors is a matter for future research.”³⁶ The Commission finds that the Postal Service has not made sufficient efforts with regard to this matter. The Commission is not persuaded by the Postal Service’s arguments to no longer investigate the issue.³⁷ In the present mail processing environment, network factors likely affect labor cost. Econometrically, such effects would depend on how the network variables are measured and how they interact with volume variables. The Commission finds it important to resolve this issue, and encourages a more in-depth investigation.

High or low R-squared values do not necessarily imply “good” or “bad” econometric models. See Stock and Watson, 2007 at 126. A high R-squared value implies that most of the variation in workhours can be predicted by the model. *Id.* The Postal Service correctly points out that the R-squared values in the Proposal Six model are high, but the “within-facility” R-squared values, which indicate how much of the variation of workhours within facilities can be explained by the models, are generally lower: 0.7076 for DBCS, 0.6357 for AFSM 100, and 0.5966 for FSS.³⁸ This finding suggests that additional variables should have been explored to improve the explanatory power of the model.

³⁵ “The overall R-squared values for the proposed labor elasticity models (equation 5 in the Variability Report)...are 0.9848 for DBCS, 0.9481 for AFSM 100, and 0.9227 for FSS.” *Id.*; see also Library Reference USPS-RM2020-13-1, folder “Analysis,” file “analysis_seasonal.txt.”

³⁶ See Docket No. R2001-1, Direct Testimony of A. Thomas Bozzo on Behalf of the United States Postal Service, September 24, 2001, at 69-70.

³⁷ See Response to CHIR No. 2, question 3.c.

³⁸ See Library Reference USPS-RM2020-13-1, folder “Analysis,” log file “analysis_seasonal.txt.”

The Commission previously stated that “variables such as the degree of support costs, space utilization, degree of flex labor, as well as several others, vary over time for the same facility and are persistently different across facilities.” PRC Op. R97-1, Appendix F at 10. The Commission further noted that “[i]f it [were] important to control for these differences in facilities over time in recovering the relationship between mail processing costs and mail volume, then...[solely using a fixed-effects procedure would be] unable to yield a valid estimate of this relationship. *Id.* The Commission suggests the Postal Service performs a thorough investigation of these issues in future variability studies.

In the future, there is also the potential for another issue with estimating a mail processing variability model. Discussing a potential impact of the COVID-19 pandemic, the Postal Service states that “[i]t may be reasonable to expect an extended period of adjustment of workhours to letter and flat distribution workloads, similar to what was observed over the Great Recession.” Response to CHIR No. 1, question 11.a. The Commission encourages the Postal Service to investigate the impact of the COVID-19 pandemic on mail processing operations in future proposals and include “recession-related control variables...to the extent that those variables would have explanatory value for mail processing workhours in recession-affected time periods.” Response to CHIR No. 4, question 4.c.

c. Monthly Dummy Variables Might Restrict Volume Variation
and Should be Replaced with Other Variables

The Proposal Six model includes dummy variables for each month to account for “residual seasonal (monthly) variability in workhours that is not explained by corresponding variations in TPF.” Variability Report at 20. Figures 8 through 10 of the Variability Report illustrate seasonal patterns in TPF and workhours across three mail processing operations subject of Proposal Six, and the Commission agrees with the Postal Service that purely seasonal effects could introduce bias into the model and therefore should be controlled for. The Postal Service contends that it uses “dummy

variables [to] estimate the effects of omitted time-varying variables.” Response to CHIR No. 6, question 5.c.; see *also* Response to CHIR No. 3, question 3.b. However, the Commission is concerned that monthly dummy variables likely control for both volume and non-volume seasonal effects, although the explanatory power from short-term variation in volumes should not be excluded from the variability model.

The Commission finds it important to reiterate its previous “tentative conclusion” that correctly specified mail processing variability models should reflect “quarterly variation” in output, but monthly dummies might restrict such variation. PRC Op. R2006-1, ¶ 3100. Discussing the “solution to the omitted variables bias” in the mail processing variability model, the Commission previously stated that “adequate statistical measures of important *non-volume-related seasonal effects* on” workhours should be included into a model as additional variables to “explicitly control for” such effects. *Id.* (emphasis added).

The problem with using monthly dummy variables is that they control for any monthly variation in the model, including both non-volume (which is desired), but also volume variation. See *id.* ¶¶ 3095-3100. As the Postal Service clearly illustrates in Proposal Six, mail processing volume for letters is the highest in December and the volume for flats is the highest in October-November.³⁹ The Commission, therefore, affirms its previous “conclusion is that quarterly variation in output should be reflected in models of mail processing labor demand.” PRC Op. R2006-1, ¶ 3100. While it is better to control for seasonal effects rather than not to control for them at all, in future variability studies, as a “solution to the omitted variables bias,” the Postal Service should “find adequate statistical measures of important non-volume related seasonal effects on work[]hours [and construct the variables to] explicitly control for” such effects. *Id.* As “the most prominent candidates,” the Commission previously identified “seasonal changes in the proportion of mail that is difficult to process (e.g., non-machinable letter

³⁹ Variability Report at 12; see *also* Library Reference USPS-RM2020-13-1, folder “Workbooks,” Excel file “Figure 4 TPF.xlsx.”

mail) and changes in the proportion of the workforce that is part- or full-time.” *Id.* Part-time and full-time employee data should be available in the TACS database, which is designed to combine all of the time keeping systems used by the Postal Service.⁴⁰ The Commission recently approved a Postal Service proposal that, among other things, separated Special Purpose Route carriers into full-time and part-time employee segments using TACS data.⁴¹

For the reasons discussed above, the Commission does not support application of the specific two-way fixed-effect model, which was considered by the Postal Service. See Response to CHIR No. 3, question 3.b. The Postal Service estimated this model with separate control variables for each month between FY 2016 and FY 2019, which allowed for more “general time-varying effects.” *Id.* The Postal Service, however, has rejected the model due to a “weak theoretical basis for non-seasonal time period effects and the statistically insignificant differences between the two-way and Proposal Six results.” *Id.* The Commission suggests that the two-way model would employ more monthly dummy variables than the Proposal Six models, which could further restrict volume variation. For that reason, and the reasons provided by the Postal Service, the Commission does not support further explorations of a two-way fixed-effects model.

D. Operational Issues

1. The Postal Service Models Should Distinguish Between Subtypes of the AFSM 100 Operations

In MODS, the AFSM 100, and DBCS operation groups distinguish between outgoing and incoming mail processing schemes, which are the different processes

⁴⁰ See TACS: Time and Attendance Collection System, Supervisor Training, Participant’s Workbook, Course 31267-01, United States Postal Service Participant’s April 2012, at 3, available at <http://inapwu.org/LeadClerkInfo/PDF-TACS%20Supervisor%20Training%20Guide%20-%202012-04-26.pdf>.

⁴¹ See Docket No. RM2021-7, Order on Analytical Principles Used in Periodic Reporting (Proposal Four), September 30, 2021, at 2 (Order No. 5991).

within an overall operation. Response to CHIR No. 2, question 5.a. The Postal Service explains that “[t]he mail processing schemes differ in the postal geography to which the mail was sorted or presorted prior to induction, and to which pieces are sorted after processing.” *Id.* The AFSM 100 operation groups distinguish between subtypes (different variations of sorting machines within an operation) of the equipment, namely the AFSM 100 machines with Automated Induction (AI) and/or the Automated Tray Handling System (ATHS), but the DBCS operation groups do not distinguish between subtypes.⁴² In the Proposal Six models, the Postal Service aligns the “variabilities with the existing labor cost pool structure in Cost Segment 3.1, which also does not distinguish subtypes [or schemes] of operations within the DBCS and AFSM 100 cost pools.” Response to CHIR No. 2, question 5.c. The Postal Service notes that it had limited options for disaggregating its models by either subtype or scheme due to the MODS data structure. Response to CHIR No. 2, question 5.a. However, the Commission suggests that it might be relevant to distinguish between sub-groups of the AFSM 100 and DBCS operations in the econometric models if there is a reason to assume that different sub-groups within these mail processing operations have different workhour variabilities.⁴³ The potential problem is that “[w]hen the behavior of economic agents [in Proposal Six, schemes and subtypes of the AFSM 100 or DBCS operations] is not the same, a regression analysis using aggregated data can provide conclusions regarding economic relationships that are different than if less aggregated data were used.” *Id.* at 1.

The Postal Service specifically discusses the modeling limitations related to two disaggregated models: (1) Incoming and Outgoing DBCS operations, and (2) AI and

⁴² *Id.* questions 5.a., 5.c. Subtypes refer to different variations of sorting machines within an operation. *Id.* question 5.a.; see also Variability Report at 6.

⁴³ For details regarding potential issues with modeling on the aggregated level, see, e.g., Thomas A. Garrett, “Aggregated vs. Disaggregated Data in Regression Analysis: Implications for Inference,” *Federal Reserve Bank of St. Louis Working Paper*, November 2002 (Garret Paper). Specifically, standard errors may be estimated incorrectly. Garret Paper at 5.

non-AI AFSM 100 operations. Response to CHIR No. 2, questions 5.a., 5.c. The Postal Service notes that “MODS operations do not distinguish subtypes of DBCS equipment,” and “flat mail preparation operations...only effectively distinguish AI from non-AI operations (not the four machine subtypes), and are not separable by scheme.”

Id. The Postal Service then estimates the variability for DBCS and AFSM 100 as a weighted average of two variabilities derived from the two relevant disaggregated models. See Table VI-7 below.

Table VI-7
DBCS and AFSM 100 Variabilities Estimated at the Aggregate Level and as a Weighted Average

| Type of Operation | Estimation Method (by the level of aggregation) | Estimated Variability | Standard Error |
|-------------------|---|-----------------------|----------------|
| DBCS | Aggregate Level (Proposal Six) ¹ | 0.976 | 0.032 |
| | Weighted Average ² of | 0.967 | 0.036 |
| | Incoming Operations | 0.956 | 0.039 |
| | Outgoing Operations | 1.092 | 0.070 |
| AFSM 100 | Aggregate Level (Proposal Six) ¹ | 0.774 | 0.091 |
| | Weighted Average ² of | 0.806 | 0.078 |
| | Mailpieces of Automatic Induction | 0.780 | 0.102 |
| | Mailpieces of Non-Automatic Induction | 0.880 | 0.071 |

Sources:

¹ Library Reference USPS-RM2020-13-1, folder "Analysis," Excel file "results_seasonal.xlsx;" log file "analysis_lag_seasonal_tests.txt."

² Response to CHIR No. 2, question 5.e.; see also Library Reference USPS-RM2020-13-3, folder "Question 5," STATA log file "analysis_seasonal_chir2_disagg.txt;" Excel file "results_seasonal_chir2_disagg.xlsx;" Excel file "CHIR 2 Q5e Table.xlsx."

For the reasons discussed below, the Commission finds that the Postal Service does not provide sufficient support for modeling the AFSM 100 mail processing operations at the aggregate level without controlling for differences between sub-types of these operations. The Postal Service itself identifies certain differences between the AI and non-AI operations. For example, the Postal Service notes that AI operations may be less flexible than non-AI operations because “[i]nteger constraints on staffing feed and prep stations can limit downward flexibility of labor usage when volumes are low and/or declining.” Variability Report at 6. The Postal Service emphasizes that it is clear that newer vintage AI AFSM 100 machines are “relatively inflexible [compared to the non-AI AFSM 100 machines] with respect to volume changes” because they require only “one clerk [to] operate all three feeders” regardless of volume. Reply Variability Report at 7. The difference between the AI and non-AI AFSM 100 workhour variabilities is 0.10, and the Postal Service states that the difference is not statistically significant. Response to CHIR No. 2, questions 5.c., 5.e. However, the Postal Service notes that the lower AI variability “lend[s] some support to the Variability Report’s observation that workhours (and hence costs) for AI equipment may be less flexible with respect to changes in workload.” Response to CHIR No. 2, question 5.c. Currently, non-automated induction operations represent 25.9 percent of the overall automated AFSM 100 workhours, a quite notable share. Response to CHIR No. 2, question 5.e. Considering the observed difference between the Proposal Six and disaggregated variabilities and the theoretical rationale for AI machinery yielding a lower variability, the Commission urges the Postal Service to consider distinguishing the AI and non-AI machines in its future variability studies.

The Postal Service also cites “practical issues” against adopting disaggregated models, but these issues are irrelevant to the operational reasons for modeling AFSM 100 on an aggregate basis. *Id.* question 5.c. The Postal Service states that “[e]rrors in workhour recording within cost pools or broader operation groups may not cancel to the same extent in a more disaggregated analysis, which may distort the calculation of accrued costs for disaggregated cost pools.” *Id.* However, as discussed in Section

VI.B.3., data cleaning must be performed at the most disaggregated level, and aggregation cannot and should not substitute for data cleaning. The Postal Service also cautions that “[d]isaggregating distribution keys would result in relatively small shares of tallies being assigned to disaggregated cost pools in some cases.” *Id.* The Commission does not endorse this argument because, as discussed, the non-AI share of total AFSM 100 workhours (25.9 percent) is not “relatively small.” *Id.* Also, as discussed in Section VI.D.2., there are other concerns related to the modeling of the AFSM 100 operations in Proposal Six, which warrant further consideration.

In contrast, the Commission finds that the Postal Service does provide sufficient support for its decision to model the DBCS mail processing operation at the aggregate level. Theoretically, there is less reason that one would expect Incoming and Outgoing DBCS operations to yield different variabilities. As the Postal Service notes, “[a]rticles in the letter mailstream tend to be more uniform in dimensions and weight, and physically more capable of high-speed processing.” Variability Report at 5 (footnote omitted). Furthermore, there is limited flexibility in staffing configuration of all DBCS machines; each machine will require one feeder and sweeper. *Id.* The Commission finds that these properties of DBCS mail processing should not vary differently between Incoming and Outgoing DBCS operations. Furthermore, the Postal Service claims that the share of Outgoing DBCS pieces has dropped from approximately 20 percent of letter automation workhours in FY 2005 to 8.4 percent for FY 2016 to FY 2019. Response to CHIR No. 2, question 5.c.

The Postal Service suggests that “[c]ombining relatively small and faster-declining outgoing operations with much larger incoming operations is also reasonable. Distinguishing them where practicable does not materially affect estimated variabilities.” Reply Variability Report at 4 (citation omitted). Empirically, the difference between the weighted average of the Incoming and Outgoing DBCS variabilities and the Proposal Six variability is less than 1 percent. See Table VI-5. As to the AFSM 100 model, the Postal Service again asserts that disaggregated modeling is not justified “given the

statistically insignificant differences between the disaggregated elasticities and the combined result.” Response to CHIR No. 2, question 5.c. The Commission agrees with the Postal Service that the small and declining share of outgoing operations,⁴⁴ and the immaterial difference between variabilities estimated in the aggregate and as the weighted average support the appropriateness of modeling of DBCS operations at the aggregate level. *Id.* Nevertheless, the Commission suggests that in future variability studies, the Postal Service continue to consider and evaluate whether it is appropriate to distinguish between subtypes of the DBCS operations when estimating labor variabilities.

2. The Postal Service Should Address the AFSM 100-Specific Issues in the Future Variability Studies

In multiple subsections of this Order, the Commission identifies and discusses issues with the Proposal Six AFSM 100 variability model. The Commission urges the Postal Service to address the issues summarized below in future variability studies.

First, the Commission finds that the AFSM 100 variabilities estimated separately for large and medium facilities are substantially lower than variabilities estimated either for small facilities or at the aggregate level in Proposal Six. See Table VI-8. This finding suggests that the Proposal Six models might not properly capture some labor-determining factors associated with facility size and volume. See *also* Section VI.C.3.b.

Second, the variability computed as a weighted average of the disaggregated AFSM 100 AI and non-AI variabilities is 0.806, which is more than 3 percentage points larger than the Proposal Six variability of 0.774. See Table VI-8. In contrast, the

⁴⁴ On average for FY 2016 to FY 2019, the share of the Outgoing DBCS operations is 8.4 percent in the DBCS workhours. Response to CHIR No. 2, questions 5.c., 5.e.; see *also* Variability Report at 4. The share of outgoing operations in the overall DBCS operations had also been consistently declining in FY 2016 through FY 2019. It was 10.0 percent in FY 2016, and declined to 7.5 percent in FY 2019. See Library Reference PRC-LR-RM2020-13-1, folder “Operational Disaggregation,” Excel file “DBCSOUT share calc.xlsx;” STATA log file “analysis_seasonal_chir2_disagg_PRC.txt.” The Commission follows the Postal Service calculation methodology of FY 2016 to FY 2019 workhours. See Library Reference USPS-RM2020-13-3, folder “Question 5,” Excel file “CHIR 2 Q5e Table.xlsx.”

difference between the weighted average variability and the Proposal Six variability for the DBCS operations is less than 1 percentage point. The implication of this difference, as well as the conceptual reasons for disaggregating AFSM 100 by AI and non-AI operations, which are discussed earlier in this section, suggest the need for serious consideration of either disaggregation or controlling for subtypes of the AFSM 100 operations within the aggregated model.

Third, as shown in Section VI.C.2., the lagged TPF variables are not jointly significant in the AFSM 100 model and the twelfth lag coefficient is not statistically significant, which indicates that the Proposal Six econometric model is not a good fit for the AFSM 100 operations.

Fourth, the Commission also finds that the estimated variabilities for AFSM 100 models change substantially when the cutoff level of productivity screening changes. See Section VI.B.4. Although the choice of the screen level is arbitrary itself, there is a clear pattern in the estimation results: the larger the screen level, the higher the variability. See Table VI-8.

Finally, the estimated variabilities for AFSM 100 operations fluctuate substantially when the sample period changes. See Table VI-8. The variability increases from 0.774 to 0.850 (or almost by 7.6 percentage points) when the sample period is reduced by one year (from 4-year to 3-year-period). *Id.* Increasing the sample period to a 5-year period results in a variability estimate of 0.731, which is 4.3 percentage points lower than the Proposal Six estimate. *Id.*; see also Section VI.C.3.b.

Table VI-8
Workhour Variability Estimates for AFSM 100 Operations

| Model Variation | Description | Variability | Standard Error |
|---|--------------------------|-------------|----------------|
| Disaggregating by Facility Size | Small | 0.901 | 0.075 |
| | Medium | 0.557 | 0.087 |
| | Large | 0.599 | 0.080 |
| Disaggregating by Subtype of the Operation | Automatic Induction | 0.780 | 0.102 |
| | Non-Automatic Induction | 0.880 | 0.071 |
| Choice of the Sample Period | 4 years, FY 2016-FY 2019 | 0.774 | 0.091 |
| | 3 years, FY 2017-FY 2019 | 0.850 | 0.085 |
| | 5 years, FY 2015-FY 2019 | 0.731 | 0.072 |
| Choice of the Productivity Screen Level | 1 percent | 0.722 | 0.095 |
| | 3 percent | 0.753 | 0.090 |
| | 5 percent | 0.774 | 0.091 |
| | 10 percent | 0.823 | 0.081 |

Sources:

Facility Size: Table VI-6; *Subtype of the Operation:* Table VI-7; *Sample Period:* Table VI-5; *Productivity Screen Level:* Table VI-3; Library Reference USPS-RM2020-13-1, folder "Analysis," Excel file "results_seasonal.xlsx."

Notes: The variabilities are statistically significant at a 5 percent level. Variabilities estimated in Proposal Six are highlighted in "green."

The Commission urges the Postal Service to investigate all the AFSM 100 issues in detail and provide appropriate revisions to the dataset and the model in the future mail processing variability studies.

E. Response to the Joint Commenters' Analysis in Support of Proposal Six

The Joint Commenters argue that empirical estimation in Proposal Six "is an improvement over the current approach of simply assuming that [mail processing costs in automated letter and flat sorting operations] are essentially 100 percent variable."

Joint Comments at 2. For the reasons discussed below, the Commission rejects this argument.

It is important to note that the current Commission methodology is not a product of “untested assumptions” as the Joint Commenters suggest. *Id.* at 1. The Commission has thoroughly considered empirical analyses conducted by the Postal Service and other parties and the expertise of operational and engineering experts when making decisions to accept and retain the current methodology for the mail processing variabilities.⁴⁵

Relatedly, not just any empirical analysis would constitute an improvement over the current methodology. In Docket No. R2006-1, the Commission reviewed three econometric models, but after considering the empirical results, it approved none of them suggesting that, “econometric modeling of mail processing labor demand variability...[had] reached an impasse.” PRC Op. R2006-1 Vol. 2, Appendix J ¶ 293. In Proposal Six, the Postal Service has made an attempt again to empirically estimate labor variabilities, but the Commission finds that the proposal suffers from multiple flaws as laid out in the current Order. As discussed in detail in this Order, Proposal Six does not provide reliable variability estimates and employs arbitrary assumptions. See Sections VI.B. through VI.D.

In support of Proposal Six, the Joint Commenters analyze the relationship between the workhours and TPF for automated flats operations in the period between FY 2012 and FY 2019.⁴⁶ The Joint Commenters find that in the analyzed time period, the total workhours in automated flats sorting operations (*i.e.*, for the combined AFSSM 100 and FSS operations) declined by less than half the [corresponding] decline in TPF. Joint Comments at 3-5; *see also* Joint Comments Workpapers. The Joint Commenters

⁴⁵ See, *e.g.*, PRC Op. R2006-1 Vol., Appendix J ¶ 1; PRC Op. R2001-1, ¶ 3024; PRC Op. R97, ¶ 3010 (citing Action of the Governors, USPS in the Matter of Postal Rate and Fee Increases, 1971, Docket No. R71-1 Decision at 4-121 to 4-129).

⁴⁶ Joint Comments at 3; *see also* Joint Comments Workpapers.

suggest that because the observed workhour reduction is much larger than “would be expected based upon USPS-estimated variabilities (approximately eighty percent) for AFSM 100 and FSS operations,” there is room for additional operational cost control. *Id.* at 4. The Commission identifies several problems with the Joint Commenters’ analysis. Overall, the analysis appears to largely ignore some of the seminal issues associated with estimating labor variabilities discussed in the CHIR responses in this docket and the Commission opinions regarding the previous variability studies.

First, and most importantly, the Joint Commenters perform illustrative analysis and do not attempt to estimate volume variability for any automated mail processing cost pool. Rather, in the Proposal Six variability analysis, the Postal Service measures a response of workhours spent in each of the three mail processing operations/machine types (DBCS, FSS and AFSM 100) to an additional mailpiece fed into a machine. Response to CHIR No. 5, question 8.a. In estimating variabilities, the Postal Service intends to capture not only “the causal effect of additional [TPF] on workhours in the [specific] operation in the same month [, but also] longer-term effects via the lagged TPF terms.” *Id.* Joint Commenters do not attempt to isolate the causal effect of volume (current or lagged) on workhours, which precludes them from providing any reliable metric for measuring the variability. Therefore, the fraction between the change in workhours and a change in volume that the Joint Commenters estimate over the FY 2012 to FY 2019 time period (and which is less than 100 percent) cannot serve as a proxy for variability in any way, and the Joint Commenters do not claim that it does.

Second, the Joint Commenters in their analyses do not account for any differences between mail processing facilities. In Proposal Six, the Postal Service uses fixed-effects to control for confounding factors that can cause differences in workhours, given a level of volume. See Section VI.C.3. For example, if the largest Postal Service facilities have layouts that cause workhours to be less than fully variable, the Joint Commenters’ analysis would not be able to isolate this effect. Joint Commenters admit that their analysis does not account for all potential confounding factors, noting that:

“[f]actors other than pure scale effects – such as management inattention to maximizing processing efficiency – likely account for some of the disconnect between workhour and TPF reduction.” Joint Comments at 4.

Third, it is inappropriate to make any propositions about variabilities for flats operations using data for 7 years from FY 2012 through FY 2019. See Joint Comments at 3-5. This is because there were operational changes during this time (such as network consolidation for DBCS and FSS AFSM 100 operations and FSS deployment), which likely affected variabilities throughout that period. See Variability Report at 11, 24. During this period, the Postal Service responded to reduced volumes by closing and consolidating mail processing facilities. *Id.* at 11. The Postal Service notes that “[t]he number of facilities with active DBCS operations declined slightly in the early PAEA period, and more significantly from FY2011-FY2013.” *Id.* The number of facilities reporting AFSM 100 activity also declined due to network consolidation beginning around FY 2013. *Id.* Figure 3. As a result, the composition of flat sorting costs does not appear to stabilize until FY 2013. *Id.* Figure 2. For FSS operations, the Postal Service notes that estimated variabilities were “relatively unstable before the deployment period beg[an] to roll out (including lags, this is in calendar year 2016 for the 48-month samples.” *Id.* at 24.

Fourth, the analysis relies on the same data as Proposal Six does, and the Commission concerns regarding data quality of workhour data apply here as well. See Section VI.B.3. The Joint Commenters have not addressed the workhour data issue in their analysis.

Finally, the Joint Commenters’ analysis appears to support the assertion “that flats sorting costs are less than fully variable.” Joint Comments at 2. For similar reasons, the Postal Service presents its runtime regressions, which use runtime as a dependent variable. See Variability Report at 22. The Postal Service states that the runtime variabilities “are a source of information based entirely on machine-collected data that bears on the validity” of labor variabilities that are less than 100 percent.

Response to CHIR No. 1, question 10. The Commission, however, bases its decision to deny Proposal Six not on the fact that the estimated variabilities are less than 100 percent, but because Proposal Six fails to provide a reliable methodology for estimating such variabilities.

The Commission concludes that the analysis by the Joint Commenters does not provide compelling evidence to warrant approval of Proposal Six.⁴⁷

VI. ORDERING PARAGRAPH

It is ordered:

For purposes of periodic reporting to the Commission, the changes in analytical principles proposed by the Postal Service in Proposal Six are denied.

By the Commission.

Erica A. Barker
Secretary

⁴⁷ The Commission does not consider the Joint Commenters' recommendation for "the Postal Service to increase passthroughs closer to more efficient levels, rather than cut [flats presort] discounts." Joint Comments at 7. This recommendation is not relevant to the current docket of updating labor variabilities for mail processing operations.